AD-A035 742 NAVAL POSTGRADUATE SCHOOL MONTEREY CALIF F/G 9/2

NTDS COMPUTER FACILITIES SCHEDULING,(U)

NOV 76 J K HARTMAN, G T HOWARD

UNCLASSIFIED NPS-55H/55H/76111 NL

1 OF 1

AD-A

035 742

END

DATE

FILMED

3-24-77

NTIS

U.S. DEPARTMENT OF COMMERCE
National Technical Information Service

AD-A035 742

NTDS COMPUTER FACILITIES SCHEDULING

NAVAL POSTGRADUATE SCHOOL
MONTEREY, CALIFORNIA

NOVEMBER 1976

NPS55Hk55Hh76111

NAVAL POSTGRADUATE SCHOOL

Monterey, California

NTDS Computer Facilities Scheduling

by

James K. Hartman
and
Gilbert T. Howard

November 1976

Approved for public release: distribution unlimited.

Prepared for:
FLTCOMDIRSYSSACT
San Diego, California

NAVAL POSTGRADUATE SCHOOL
Monterey, California

Rear Admiral Isham Linder
Superintendent

Jack R. Borsting
Provost

This work was supported by FLTCOMDIRSVSSACT, San Diego.

This report was prepared by:

James K Hartman

James K. Hartman, Assoc. Professor
Department of Operations Research

Gilbert T. Howard

Gilbert T. Howard, Assoc. Professor
Department of Operations Research

Reviewed by:

Michael G. Sovereign

Michael G. Sovereign, Chairman
Department of Operations Research

Robert Fossum
Dean of Research

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER NPS55HH55HH76111	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) NTDS Computer Facilities Scheduling		5. TYPE OF REPORT & PERIOD COVERED
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) James K. Hartman Gilbert T. Howard		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Naval Postgraduate School Monterey, California 93940		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS FLTCOMDIRSYSSACT San Diego, California		12. REPORT DATE November 1976
		13. NUMBER OF PAGES ~~44~~ 42
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for Public release: distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This report investigates the scheduling of NTDS computer facilities at FCDSSA, San Diego. NTDS mock-up, digital computers, and various items of computer peripheral equipment are combined in many different configurations for use in training and system test by FCDSSA. Several systems for scheduling the users on this equipment are proposed and evaluated.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

ABSTRACT

This report investigates the scheduling of NTDS computer facilities at FCDSSA, San Diego. NTDS mock-up, digital computers, and various items of computer peripheral equipment are combined in many different configurations for use in training and system test by FCDSSA. Several systems for scheduling the users on this equipment are proposed and evaluated.

TABLE OF CONTENTS

I. INTRODUCTION AND BACKGROUND

A. Purposes

This report investigates the scheduling of Naval Tactical Data System (NTDS) mockups and the associated computer facilities at the Fleet Combat Direction System Support Activity (FCDSSA) and the Fleet Combat Direction System Training Center, Pacific (FCDSTCP), San Diego. NTDS mockups, digital computers, and various items of computer peripheral equipment are combined into many different configurations for use in training by FCDSTCP and for use in program development and program test by FCDSSA. Equipment must also be made available periodically for required maintenance. The variety of different users and configurations creates a problem in job and equipment scheduling which is quite complicated. The purpose of this report is to investigate possible systems for scheduling, and to consider the feasibility and the desirability of automating the scheduling task, a task which is now performed manually.

B. Concerns with Manual Scheduling

Initial discussions indicated some areas of concern with manual scheduling. The most obvious is that the quality of schedule produced depends on the skill of the individual scheduler, and that the scheduling process is not formalized enough to readily teach it to a new scheduler. A second problem is that the individual primarily responsible for scheduling is only on duty during the day shift. Equipment failures frequently require schedule revisions at other times, when the absence of the scheduler may lead to some confusion and cancellation of jobs when minor adjustments would permit them to run. Scheduling is an inherently complex task with many possible alternatives for each of the many choices to be made. Manual schedulers can explicitly consider only a few of these alternatives. Probably there is more information available in the system than the manual scheduler can use. The speed and information processing capacity of

the digital computer suggest that an automated scheduling system can consider more information and more alternatives, and hence may be able to produce schedules which are better.

C. Organization of This Report

Section II of this report will discuss the performance requirements that should be imposed on a scheduling system, and the data requirements for such a system (whether manual or automated). In Section III we discuss concepts of job priorities and their implications for scheduling systems design. Section IV presents several design issues and measures of effectiveness for analyzing and comparing alternative systems. In Section V five alternative scheduling systems are described and analyzed with regard to efficiency of the schedules produced, flexibility, organizational context and impact on the user, and feasibility of computerized implementation. These comparisons highlight some important issues in scheduling system design and lead in Section VI to recommendations and to choices that should be made prior to detailed design and implementation of any automated system.

II. PERFORMANCE REQUIREMENTS AND DATA REQUIREMENTS

In this section we discuss the facilities scheduling problem from the standpoint of the user, emphasizing the services which a scheduling system should provide for the user and the data which the user provides for the system.

A. Performance Requirements

The basic function of a scheduling system is to accept a job request from the user and to respond with an assignment of time and equipment which will allow the job to be performed. To the extent that it is possible the time assigned should be convenient to the user and the equipment assigned should form a configuration which is spatially convenient and electronically compatible. The schedule should be provided early enough to allow adequate time for preparation and planning by the users.

Once a schedule is published, a second basic function of the scheduling system is to keep the schedule up to date. Changes in the schedule may be required for a variety of reasons. User requirements may change because of unforeseen problems. New jobs may arrive to be scheduled, and sometimes these jobs are so important that they have a priority status--the schedule must be changed to accommodate them. Changes in the maintenance status of equipment (equipment failures) may result in decreased equipment availability, and if this equipment is currently scheduled to be used, the schedule will require modification. In addition to creating a modified schedule, the scheduling system must communicate the changes to any users who are affected. If possible the scheduling system should operate in a way that minimizes the occurrence of schedule changes.

In addition to the basic functions of creating and updating the schedule, a scheduling system can also provide a variety of summary reports

about the NTDS equipment and its utilization. The data processing capabilities of a computerized system would be particularly useful for such management reporting.

B. Flexibility

For a scheduling system to be useful to a variety of users with differing needs, it must be flexible enough to deal with those needs. The system should be able to function with minimum input, but should be able to handle additional data (e.g. if a certain user desires not to be scheduled on Wednesday, the system should be able to accomodate his needs). It should be able to use priority information in the scheduling process, but also, if priorities are not assigned, should be able to schedule without them. It might be desirable for the scheduler to offer alternatives to users if their first preferences cannot be met. The system should be able to modify the schedule at any time around-the-clock especially when significant amounts of work are being done during the night shifts. To the extent that it is possible the schedule should reflect user preferences for working hours and days, but in periods of high demand the system must be capable of scheduling jobs in spite of expressed preferences. Finally, the system must allow for a manual override in the (hopefully) rare situations which it cannot handle routinely.

All of this flexibility is purchased at the cost of increasing system complexity. Substantial care should be exercised in the design stages to ensure adequate flexibility at reasonable cost.

C. Data Considerations

The major data elements that will be involved in a scheduling system can be categorized as input data, output reports, and continuing data base entries.

1. Input Requirements

a. Work Requests

The primary input to a scheduling system is a user-generated work request. This input must include job or user identification, a list of the equipment required, and job duration, and may also include priority information and time preferences if the user desires to include them. We postpone consideration of how or when requests are submitted until the discussion of individual systems in section V.

b. Maintenance Status Changes

If the scheduling system is to be responsive to equipment availability both in the initial scheduling and in schedule revision, it must be aware of the maintenance status of all equipment. Changes in maintenance status would be input by authorized maintenance personnel. Interfaces with the existing Equipment Status Program should be explored.

2. Output Produced

A wide variety of outputs might be produced by a scheduling system. Some will occur in hard copy (paper), while others might be better produced in soft copy (phone call, video display unit). At this time we list the most important outputs postponing discussion of timing and formats to section V.

a. Master Schedule

The primay output of a scheduling system is a master schedule which gives an up-to-date list of the scheduled time for all currently scheduled jobs. The master schedule must be available to all users of the system.

b. Schedule Detail

On request the scheduling system should give detailed information

about any scheduled job (such as the precise pieces of equipment assigned to that job). In some systems this detail might be included as a part of the master schedule.

c. Change Reporting

When changes in equipment maintenance status or arrival of unexpected high priority jobs force a change in the current schedule, any job which is affected must be notified of the change (cancellation, new time, different equipment). Such changes will also appear in the revised master schedule which should be kept up to date at all times. On demand the system might also publish information about previously scheduled equipment which becomes available because of a schedule change.

d. Summary Statistics

A variety of summary statistics and management reports can be prepared form the basic data which the scheduler processes. As a sample we list equipment utilization, equipment reliability statistics, total system load and system bottlenecks. Such reports may be useful for monitoring the training and program test operations, for recognizing problems in FCDSSA operations, and for aiding decisions such as which additional equipment to buy (if any).

3. Data Base

A scheduling system must have access to certain data to perform the scheduling task. For a human scheduler much of this data may be carried in memory; for a computer scheduling system various files must be established. It is convenient to categorize the data items as follows:

a. Permanent Files

Permanent files are those which rarely change. In permanent files the system must maintain a list of all the equipment to be scheduled

along with sufficient information to determine which units are electronically compatible and which combine to yield convenient workspace layouts.

b. Semi Permanent Files

In Semi Permanent Files the basic content is stable but periodic changes occur to details. An example is the current maintenance status for each piece of equipment the scheduler processes. It might also be useful for the system to remember regularly occurring schedule requests such as for regularly scheduled maintenance or recurrent training.

c. Transient Files

Transient files change totally from week to week. Data items which would be included in such files are the current master schedule with its supporting detail and a list of pending requests (if any).

d. Summary Files

Production of summary statistics and management reports may require saving historical schedule data past the time when it is current. The required files will, of course, depend on the reports desired.

III. PRIORITY CONCEPTS AND IMPLICATIONS

A cursory examination of the current FCDSSA manual scheduling process makes it clear that job requests do not all receive the same treatment. For example, some allocations of time and equipment in the schedule are mandatory, such as regularly scheduled periods for routine maintenance. During the day shift, the FCDSTCP training activities have priority over FCDSSA program development and test activities. Some program development and test job requests are more important or more critical than others (such as a job request from Test and Delivery for an important NTDS program which is approaching its installation date), and the scheduler is instructed to give them preferential scheduling treatment.

These examples suggest that in any effective scheduling system there must be some means for dealing with priority relationships among the various job requests. The extent to which priorities are established and used is a management decision, but the system should be capable of handling whatever priorities are assigned. Thus at one extreme we might imagine a system in which no priorities are assigned and in which all jobs are treated equally, while at the other extreme every job might have a unique number assigned which measures its ranking in a priority hierarchy. The present system lies between these two extremes: jobs are grouped into several classes, with priorities among classes implicitly determined by the rules under which the scheduler operates. Within each class all jobs are treated equally.

The key issue in discussing priorities is not the assignment of priority numbers, but rather the establishment of clear, well defined, and agreed upon rules that explain precisely what it means for one job to have priority over another. In this section we discuss several concepts of job priority and indicate some implications for scheduling system design.

A. Reasons for Priorities

The reason for having priorities differs from case to case. For important program development and test jobs which are approaching a firm deadline, the reason is clear -- it really is more important and more critical for these jobs to be run than other jobs with longer lead times. For training jobs during the day, and for routine scheduled maintenance, it is not at all clear that every one of these jobs is more important than any of the other jobs over which they have priority. In the aggregate, however, training and maintenance are jobs which must be done on a continuing basis, and giving them priority during certain times is a convenient way of allocating time and equipment between these jobs and other nonrecurrent jobs. Thus, a priority system must be able to distinguish between jobs which are important on a continuing basis and jobs which occasionally become super critical, and the system must have rules which let it choose between such jobs if they compete for resources.

B. Uses of Priorities

There are several ways in which priorities may be used within a scheduling system.

Sometimes priorities will be the cause of a schedule revision. On occasion job requests arrive which are so important that they must be met immediately even if jobs already on the schedule must be moved or cancelled to accommodate the new request. We say that such jobs have a pre-emptive priority. It is clear that pre-emptive priority jobs can severely disrupt normal operations. Care should be taken that premptive status is only assigned to jobs which genuinely require this special priority.

It should be noted that the scheduling time horizon interacts with the notion of pre-emption, since any job that can be anticipated far enough

in advance can be scheduled in the regular scheduling process, and thus will not need to pre-empt anyone else already scheduled. Only important jobs which cannot be anticipated in time for the normal scheduling cycle need to have pre-emptive priority assigned. We would expect this to happen more frequently in systems where the time horizon is relatively longer.

Another use of priorities is in the initial preparation of the master schedule. There are generally many job requests waiting to be scheduled, and the scheduling can be done in such a way that the needs of the higher priority jobs are satisfied before those of lower priority jobs. If the total resources demanded by all job requests exceed the available resources, then lower priority jobs are postponed to later dates.

This use of priority rankings might be called <u>preferential priority</u>. It is clearly a weaker concept than pre-emptive priority, and rather than disrupting the schedule, can be a considerable aid in formulating a good schedule.

When schedule revisions are necessary (perhaps due to equipment failures or arrival of pre-emptive jobs) it may be necessary to cancel a job which is already on the schedule. Generally there will be several jobs which could be cancelled to make needed equipment available, and of these the lowest priority job might be cancelled with its equipment reallocated to replace the malfunctioning equipment. If such cancellations are necessary, priority information can help to minimize the impact of the cancellation.

C. Implications of Priorities

Adherence to priority rankings will restrict the possible choices open to a scheduler whether human or automated. This restriction is desirable because it forces the schedule to reflect the importance of the various job requests. Such restriction may also, however, have adverse effects.

A schedule which follows priority rules may make less efficient use of available equipment because it has fewer possible alternatives to choose from than a schedule without the added priority constraints. Thus, to retain maximum flexibility for the scheduler, and to attain efficient equipment utilization, the priority restrictions imposed should be the minimum necessary to accurately reflect relative job importance. A system with several broad classes of priority would be better in this regard than a system in which each user had a unique priority different from all others.

Multiplication of priorities beyond those required also has the undesirable effect from a systems design viewpoint that the system must have clear and unequivocal rules for dealing with priority information in the scheduling process. Overly complex systems will be difficult to specify, hard to implement, and impossible to explain to the user. Systems that users cannot understand are not likely to be successful.

D. A Proposed Priority System

The following job priority structure is proposed for incorporation into a scheduling system for FCDSSA/FCDSTCP. Job requests are divided into four broad classes: Training, Maintenance, Pre-emptive, and Regular. Within each class numerical preference priority levels could be established by appropriate authority where required by differing job importance. The number of distinct levels in any class should be kept as small as possible (perhaps 2 or 3 would suffice). Training and Maintenance jobs would have higher priority than Regular jobs in scheduling during certain designated times, but lower priority otherwise (as is the case in the current system). Only Pre-emptive jobs would be able to displace other jobs in an existing schedule. Jobs would be classified as pre-emptive only by command authority with the implication that these

jobs are rather rare. Within each of the four classes, the numerical levels determine preference treatment in scheduling and schedule revision.

Details of the interactions of such priorities with other aspects of the scheduling system are reserved for Section V, where complete systems descriptions are given and analyzed.

It is believed that this proposed priority structure is flexible enough to accurately reflect the meanings and uses of relative job importance in the scheduling system, and also simple enough to be easily understood and easily incorporated into the decision logic of either a manual or an automated job scheduling system.

IV. ISSUES FOR ANALYZING AND COMPARING ALTERNATIVE SCHEDULING SYSTEMS

As a basis for comparing possible alternative scheduling systems, we present in this section a discussion of some measures of effectiveness for system design. It should be clear from the outset that in a problem area with as many interactions as the FCDSSA scheduling problem, there are many impacts on individuals and on the organization and thus many potential measures of the system's effectiveness. Some of these will be difficult or impossible to quantify and in the final system design it will be essential to consider interactions and tradeoffs among several measures. The primary tradeoff is between cost and effectiveness. At this preliminary stage, cost is difficult to assess, so our analysis will concentrate primarily on system effectiveness.

A. Issues for Analysis of Alternative Systems

Seven issues in scheduling system design will serve as the framework for analysis and comparison in Section V, so we briefly present these issues here.

1. Schedule Optimization

Under this heading we will discuss the number of alternatives available to the scheduling system. A system which has more alternative choices available should be able to schedule jobs in a more efficient manner, thus leading to a schedule which is closer to optimal. It also follows that a scheduler which is expected to do more optimization will require more intricate decision logic to make the required choices between alternatives.

2. Priorities

An important measure of effectiveness for scheduling system design is the extent to which the system can use priority information in the scheduling process.

3. <u>Timing</u>

We will consider under "Timing" issues such as when and how often the schedule is produced and updated, when job requests can be input into the system, the maximum, minimum, and average lead time (difference between the time the job request is received and the time the job actually starts), and the delay to be anticipated if a user's job must be cancelled and rescheduled.

4. <u>Stability</u>

All schedules, once prepared, are subject to change for a variety of reasons. Stability refers to the extent to which scheduled jobs are protected against changes.

5. <u>Effect of Cancellations</u>

If jobs must be cancelled from an existing schedule it is necessary to decide which particular job (or jobs) to cancel and what to do with the cancelled job to get it rescheduled.

6. <u>Effects on the User</u>

A variety of questions are considered under this heading. Is the system convenient to use and easy to understand, or might it be confusing? Can the system make use of expressed user preferences? Is there opportunity for involving the user in the scheduling process in an interactive manner?

7. <u>Computer Facilities Required</u>

At this preliminary stage of system definition it is difficult to estimate the precise magnitude of the computation and storage required if a system is to be automated. We can, however, provide initial information about frequency of system usage, relative magnitude of the computational tasks, and type of system access required.

B. <u>General Issues</u>

Some more general dimensions of system design, which will not be considered separately for each alternate system, but which should be kept in mind

during the design process, are presented in the remainder of this section.

1. Centralization - Decentralization

Manual systems tend to be centralized in the sense that there is an individual, the scheduler, who handles all job requests, does the scheduling, and communicates with the users (including maintenance). The scheduler is a focal point to whom management can go for information about the system or to implement changes. Automated systems may be designed to be centralized with a designated scheduler maintaining the role of communication with both the computer and the user. Alternatively in an automated system the role of scheduler may be abolished and individual users may communicate directly with the scheduling system using remote terminals. The issues of management information and control in the decentralized system should be considered before such a system is adopted.

2. Degree of Automation

Scheduling systems can range from fully manual to computer assisted to fully automated. The scheduling system consists of several functions, each of which can be performed in a variety of ways. These functions include the decision making function in which jobs are assigned scheduled times, the data storage and retrieval function which deals with such information as equipment status, user requests, and job priorities, and the function of communication between the system and the users. Each of these functions can be done manually or by a computer. Hybrid systems in which some functions are automated and others are done manually should be considered as well as the extremes.

3. Failure

Whatever system is used, consideration must be given to how the scheduling will be done in the event of a system failure. In a manual system "failure" can be thirty days leave or illness for the scheduler. In an automated

system "failure" might be a computer problem. Care should be taken that a hardcopy of the current schedule is always available. This would probably be adequate planning since it is unlikely that a computer failure in the scheduling system would not be repaired in a few days, but still one should consider what may happen if full reliance is ultimately placed in an automated system.

4. Schedule Efficiency and Extreme Conditions

When the total of all job requests is moderate so that there is substantial excess equipment capacity, almost any scheduling system should be able to develop workable schedules. It is doubtful whether "efficiency" has any meaning in this context. In periods of high demand, efficient scheduling may increase the total number of jobs that can be run, so that in these circumstances some degree of efficiency is desirable. Any proposed system should be tested under extreme conditions such as heavy total load and short deadline times, for it is in these conditions that the differences between systems will become apparent. The impact of unusually frequent equipment failures is another extreme condition that might be investigated.

V. ANALYSIS OF SOME ALTERNATIVE SCHEDULING SYSTEMS

In this section we present five alternative system designs for the FCDSSA facilities scheduling problem. By presenting some concrete alternatives, we can illustrate the interactions between system design and measures of effectiveness.

A. Weekly Batch System

In the weekly batch system the job requests are accepted at any time up to some deadline, but not processed until that deadline is reached. Once processing is begun, all currently available requests are considered in producing the schedule for the next week. Any jobs which cannot be scheduled for the coming period are deferred to the following period. Any job cancelled because of equipment failure will also be deferred to the following period, unless it is possible to reschedule it in a slack period in the existing schedule.

In the scheduling process a priority system is used to assign the times. Normally a job once scheduled will not be cancelled except when a pre-emptive priority job arrives or when equipment failure demands it. Cancellation is also done on a priority basis.

The schedule for the next period would be available so that users could submit additional requests during the period for specific times and equipment not already assigned. Thus each arriving request must be screened to determine if it is for the current or the coming period. Real time cancellations and rescheduling would be handled by that portion of the system which screens requests.

The Weekly Batch system is closest to the scheduling system which now exists in manual form at FCDSSA. In this discussion we will assume that this system has been made available round the clock for flexibility in rescheduling

when changes occur (perhaps by developing a computerized version of the weekly batch system), and that a job priority system has been implemented to aid in assessing relative job importance during the scheduling and rescheduling operations.

1. Schedule Optimization

Because it accumulates an entire week's job requests before doing any scheduling, the batch system has the most alternatives to consider when doing the scheduling task. It follows that it should be able to produce schedules which make the most efficient use of the available facilities. Because of the large number of alternatives, it also follows that detailed design of the scheduling decision logic for a computerized batch system will be more difficult than for the other systems presented.

2. Priorities

The simultaneous processing of an entire week's job requests allows more priority relationships to be considered and acted upon in the scheduling process than in systems which schedule fewer jobs at one time. Priorities are also available to assist in selecting jobs to be cancelled if it should be necessary.

3. Timing

In the weekly batch system a schedule is created once a week to include all the job requests for the following week. To use this system effectively a user must be able to anticipate his job needs by about a week. For such a user, and especially for users with continuing requests from week to week, a weekly scheduling deadline is quite convenient. A user whose jobs are not so predictable may, however, have problems, since a job need which arises suddenly on Friday of this week (after the weekly scheduling deadline) will not be scheduled for next week (unless it happens to fit an open space in the

schedule). When it is finally scheduled for the second week hence, its scheduled time might be as late as Friday, leaving a maximum of a two week delay between the initial request and the actual run time. The other side of this coin is that a user who submits his request just before this week's schedule deadline may be scheduled for early next Monday, leaving a rather short time horizon for preparation and planning. The delay between job request submission and actual run time is, thus, variable (3 days to 2 weeks) depending on when in the week the request is submitted. This variability will be of no consequence to some users, but might be a problem for others.

4. Stability

In the weekly batch system, once a job request is scheduled, the time and equipment allocated to it are considered firm unless equipment failures or preemptive jobs force a change. Thus the user normally need not be concerned that his scheduled time will be changed.

5. Effect of Cancellations

When a pre-emptive priority job or an equipment failure occurs, if there is no way to adjust the allocation of equipment to meet all the requirements, then some job (or jobs) must be cancelled. Among all the jobs whose cancellation would restore the schedule to feasibility, cancellations should occur in priority sequence. The cancelled job must then be rescheduled. Here the long system lead time may again be a problem if the rescheduling cannot be done in the current week's schedule.

6. Effects on the User

The Weekly Batch system should be an easy system to use. Its similarlity to the current FCDSSA manual scheduling system means that, from the user's point of view, changeover and implementation problems should be minimal. It is easy to understand. The simultaneous processing of an entire

week's requests means that the system should be able to make good use of the preferences expressed by the users and of their relative priorities. In a Weekly Batch system, jobs are held for up to a week before being scheduled. Then all the scheduling occurs within a short scheduling period. This time structure implies that it would be cumbersome to try to involve the user interactively in the scheduling process.

7. Computer Facilities Required

If the weekly batch system is computerized, it will require one rather extensive scheduling program execution per week to do the initial weekly scheduling. Schedule changes during the week and accumulation of new job requests for the next week could also be automated and would require relatively frequent running (several times per day) of rather small programs.

8. Advantages and Disadvantages

In summary, the major advantages of a Weekly Batch system, in addition to its familiarity to current users, is the degree of optimization which can be exercised in the initial scheduling process because of the large number of jobs being processed simultaneously. The major disadvantage is the inherent variability in schedule lead times with the potential of long delays for some users. It would seem that there is a tradeoff in system design between efficiency in overall usage of the NTDS equipment which is being scheduled and efficiency in time utilization for the individual users. The Weekly Batch system is biased towards efficient equipment utilization.

B. Continuous System

In this case, the scheduling system is assumed to operate continuously. Job requests are received at any time and processed immediately. Each newly arriving request is assigned a scheduled time from among those available. Once scheduled, a job is not removed from the schedule except for the arrival of a job with pre-emptive priority or because of equipment failure. If removed, a job simply re-enters the system as a new request. Cancellation is done by considering the priority of existing jobs in the schedule and cancelling those of lowest priority.

The schedule is always available and users can scan it to determine when equipment is available and submit a request for that time.

1. Schedule Optimization

Among the systems presented in this report the continuous system is unique in the fact that no optimization is performed by the system in assigning times to job requests. Since each request is processed as received and is permanently assigned a time and a set of equipment, no opportunity exists for planning and selecting among alternative schedules. The mechanics of the algorithm for implementing this type of system would be very simple. All that is required is to search the existing schedule for feasible times. Among those times the one which most nearly matches the user's expressed preference would be selected.

2. Priorities

In the continuous system, priorities cannot be used in the initial scheduling process since each request is processed upon receipt and assigned a time. If cancellation is required because of the arrival of a pre-emptive priority job, cancellation can be done by considering priorities, although there would be very few options if the pre-emptive job designated the time and equipment needed.

3. Timing

The continuous system would accept job requests at any time and would thus appear convenient to the user. However, because of the inability of the system to arrange the schedule efficiently, it is almost certain that under moderate or heavy loads the delay until the job is run would be quite large.

4. Stability

No changes are made in the schedule under the continuous system except when a cancellation is required.

5. Effects of Cancellation

Job cancellation would occur only when a pre-emptive priority job arrives or when equipment fails. If the cancellation is caused by a pre-emptive job, some options might be available regarding which job or jobs to cancel. If so, the decision could be made by considering user priorities. Alternatively, the arrival of a pre-emptive job could be treated exactly like a machine failure. The arriving job would simply declare that during the required time certain pieces of equipment are "out of service." The users previously scheduled on that equipment would be the ones cancelled. A cancelled job would simply re-enter the system as a new request.

6. Effects on The User

This system offers the user the opportunity to interact with the scheduler in selecting a time assignment. Once scheduled, the user would know that his assignment is secure. There is little reason for user confusion in the use of this system, but several reasons for user dissatisfaction. Among them are the possible long delay (due to inefficiency in the schedules) between submitting the request and running the job, and the fact that when cancelled, a user simply returns to the end of the line.

7. Computer Facilities Required

To implement the continuous scheduling system using a computer would require use of an interactive system which is available at all times. A relatively small amount of computation would be required since this system is unsophisticated. To emphasize this point, imagine that the system was implemented without a computer. It would be adequate to maintain a scheduling board where each user "helps himself" by reserving the required equipment and times on a first-come-first-served basis. The only storage required is to maintain the current schedule and the equipment status (for further scheduling).

8. Advantages and Disadvantages

The major advantages of this system are in the simplicity of implementation and in the apparent convenience it offers to the user in accepting and processing his requests interactively and immediately. This apparent advantage is probably outweighed by the inefficient usage of the equipment and time that is inherent in this system.

In summary, the system is definitely biased away from efficient equipment utilization, and the advantages to the user are not sufficient to compensate for this loss. This, coupled with the fact the system makes almost no use of priority information indicates that the system is not adequate for FCDSSA.

C. Batch/Continuous System

Some of the features of the batch system and some of the continuous system are combined in this system. Here a portion of the equipment is scheduled on a periodic basis (say, 1 week) while the remainder of the equipment is intentionally left for scheduling on a daily basis. The weekly scheduling considers all requests received up until that time which are designated for the weekly system. These are considered on a priority basis in establishing the week's schedule. (If the workload is high, a larger portion of the facilities may be allocated to the weekly scheduling system.) Requests received during the week which are designated for the daily system are processed immediately and assigned a time during the current week, if available.

Jobs using the weekly scheduling system are not cancelled except when equipment failure demands it or a pre-emptive priority job arrives. When a cancellation is required,those jobs assigned under the daily system are cancelled before those assigned under the weekly system.

The Batch/Continuous system is an attempt to combine the long lead time and schedule optimization of the batch system for users who can profitably use it, with the short term convenience of the continuous system for users who require this flexibility. This is done by reserving a portion of the facilities for continuous users. For concreteness in the following discussion let us arbitrarily say that 40% of the facilities are reserved for continuous users. Then essentially there are two systems operating in parallel: a batch system with 60% of the equipment to be scheduled and a continuous system with the remaining 40%. We have already analyzed batch and continuous systems, and the characteristics of the mixed system will correspond to those of the batch system for 60% of the equipment and to those of the continuous system for 40% of the equipment. We will not repeat these analyses here, but rather focus on other aspects of the system.

There is one important decision to be made in this system -- the proportion of the facilities to be reserved for the batch mode versus the continuous mode. This proportion can either be fixed and held constant, or allowed to vary from week to week depending on the demand for batch services.

If the proportion allocated to batch jobs is allowed to change from week to week to meet the pending batch job requests, then this system is essentially a pure batch system -- the only facilities that are "reserved" for continuous job requests are the facilities that were left over when all batch job requests were satisfied. In times of high demand, continuous users might well find that they could not run since essentially no facilities were left for them after batch jobs were scheduled.

If the proportions of the facilities allocated to batch and continuous are fixed, then inefficient schedules may result if the actual job request submissions vary from this proportion. For example, suppose the allocation is 60% - 40% as indicated earlier, and suppose 80% of the facilities are requested by batch jobs in a particular week. Then the system with a fixed 60% - 40% allocation would schedule in batch mode up to the 60% limit, and there would be 20% of the batch job requests left unscheduled. Something must be done with this 20%, and the only reasonable thing to do seems to be to immediately schedule them via the continuous mode. When this was done, the remaining equipment would be available for continuous requests arriving throughout the rest of the week.

It seems clear that nothing favorable has been accomplished by this procedure. Twenty per cent of the jobs have been treated as continuous instead of the batch mode they requested The scheduler has been forced to consider an 80% job request block in two pieces -- 60% first and then the remaining 20%, and this probably will result in a less efficient schedule than if all

80% had been scheduled simultaneously in batch mode. The only system facilities reserved for the truly continuous requests arriving during the next week are the leftovers from scheduling the initial 80% requests, and the suboptimal scheduling of that 80% implies that less may be left over.

In summary, this Batch/Continuous system does not provide anything for either class of user that is not readily available in the pure batch system in which left-over equipment is scheduled to any arriving job request on a continuing basis throughout the week.

In the above analysis we have purposely oversimplified. It is clear that in describing the resources required by a job request or reserved for a class of job requests, it is inadequate to merely state a precentage level, since there are many different kinds of equipment to be scheduled. The batch job requests for a given week might require 100% of the available capacity of a particular mockup, 80% of the computer availability, 50% of the magnetic tape units, and 0% of a particular simulator that is not involved in any of the requests. Allowing for this diversity of system facilities, only compounds the difficulties mentioned above and increases the potential degree of suboptimality in the schedules produced.

D. Continuous, Initially Tentative System (CIT System)

The CIT system operates continuously, accepting job requests at any time, and processing each request as received. A scheduled time is assigned to each job request considering its priority, expressed preferences, and the existing schedule. Each assignment is flagged as tentative or permanent. Those jobs scheduled beyond some horizon (say 3 days) are labeled as tentative, others as permanent. Tentative assignments automatically become permanent as time advances, so that the scheduled time lies within the 3-day horizon. Tentatively scheduled jobs are subject to rescheduling in the event that a higher priority job enters the system with a request that cannot otherwise be satisfied. To reduce the amount of rescheduling, permanently scheduled jobs are never cancelled except when equipment failure necessitates it or when a job with preemptive priority requires it. In the event that some permanent or tentative job must be cancelled, the decision of which one to cancel is made by considering the priority level of the existing jobs in the schedule and cancelling that set of jobs with the lowest priority. Thus two low priority jobs might be cancelled rather than one of higher priority.

Any time which becomes available within the permanent horizon is automatically filled by the scheduling system if possible by considering the set of tentatively scheduled jobs and moving any of those, on a priority basis, into the available time if the job request can be satisfied by the move. Thus user preferences would be considered in making such a move. Any user who, upon viewing the current schedule, finds a time within the permanent horizon which is useful to him, can request this time immediately by simply submitting a job request for it.

1. Schedule Optimization

This system can be viewed as a modification of the continuous system in that it receives requests at any time and processes them immediately.

In this system each user is assigned a tentative or permanent time with the realization that tentative times are subject to change as the scheduling system receives further requests. One reason for possible changes is that through such changes more efficient use of the equipment might be made. Thus a limited amount of optimization is done in this system and to that extent it is more efficient than the continuous system.

2. Priorities

As contrasted with the continuous system, the CIT system is able to make use of priority information. In the initial scheduling process tentative times are assigned on a first-come first-served basis after considering user preferences. While it is flagged as tentative, a job is subject to rescheduling if a conflicting higher priority job arrives. In job cancellation the use of priorities parallels the continuous system.

3. Timing

There is no delay between the initial job submission and the receipt of a scheduled time, but the scheduled time may be any number of days in the future. If the job time is 3 or more days away so that it is marked tentative, the user is still unable to plan his work with the assurance that the scheduled time will become permanent. It is possible that his job will be moved either way, causing him either the inconvenience of rushing to prepare for it or the annoyance of being delayed.

4. Stability

This system handles the issue of the schedule's stability in a unique way, using the tentative-permanent designations. The intent is to prohibit schedule changes in the near term (i.e. within the permanent horizon). Stability is not guaranteed beyond that time and it is possible that considerable rescheduling will occur beyond the horizon.

5. Effect of Cancellation

The "cancellation" of a tentative job is simply the assignment of a new tentative time to the job. The cause for such a change could be the arrival of a high priority job or the need to rearrange the schedule to obtain more efficient use of the resources in view of the current set of tentative jobs. Depending on the specific system design, reassignment might be done for a reason which is not directly related to the user cancelled. For example, it may be efficient to move a tentatively scheduled job from Thursday to Friday because a newly arriving job fits better on Thursday. A variety of decision rules is possible in resolving these problems. If a permanently scheduled job must be cancelled, the decision of which to cancel could involve the job priorities.

6. Effects on the User

This system offers security to the user's scheduled time within the permanent horizon, but offers no security beyond that. Users would almost certainly experience difficulty in planning their work beyond the permanent horizon and would possibly soon be forced to ignore tentative assignments until they became permanent. Without a detailed explanation of why tentative assignments are changed, the system would appear confusing from the user's point of view. The opportunity for user interaction in scheduling is more apparent than real, since the system might change any initially selected time. without further consulting the user, although his expressed preferences could be considered in making such a change.

7. Computer Facilities Required

This system would require that a terminal (or terminals) be available continuously for receiving job requests. The system would have to interact with a scheduling algorithm for selecting an initial assignment. The amount of computation in the algorithm would be less than in the weekly batch

system, but more than in the continuous system. The system would need to interact with files storing the current schedule, the equipment status, and perhaps user preferences in addition to a file of compatible equipment configuration. In contrast, the continuous system need not store user preferences nor access a file of compatible equipment since in that system the user could view what is already assigned and simply request any remaining equipment desired.

8. <u>Advantages and Disadvantages</u>

One advantage of this system is its ability to operate in a continuous mode but to also do a certain amount of optimization on the schedule produced. The primary disadvantage is probably in the lack of security offered to the user for tentatively scheduled jobs.

E. Daily Scheduling with 3-Day Horizon

This scheduling method accepts job requests at any time and checks to see if the job can be run in the current day's schedule. If not, the system stockpiles the request in a pool of active requests. At the end of each day (day 1), this pool is processed to produce the schedule for the 4th day hence. During this time jobs are also considered for processing in the existing schedule for days 2 and 3. Jobs are considered in priority order and individual time preferences are also permitted. After the schedule for day 4 is produced, all scheduled jobs are assumed to be permanently scheduled and will be cancelled only if necessitated by equipment failure or by the arrival of a job with preemptive priority. If a job must be cancelled, the job priorities are considered in selecting the job or jobs to cancel.

If a job is cancelled or withdrawn so that some equipment becomes available, the system checks the pool of active requests to determine if that equipment can be used by some waiting job. This check is necessitated by any change which might occur in the schedule between regular scheduling periods.

In this system the schedule exists for the current day as well as for days 2 and 3. Users can check the schedule to determine the availability of equipment and can submit a job request for the current day if equipment is available.

1. Schedule Optimization

Schedule Optimization in this system is at an intermediate level in that an entire day's schedule is produced in each scheduling cycle. Less opportunity exists for optimization here than in the weekly batch system, but more optimization is done here than in the continuous or continuous-initially-tentative systems.

2. Priorities

The daily scheduling system is able to make use of priority information in the initial scheduling process as well as in job cancellation. Initially jobs are scheduled from the pool into days 2, 3, and 4 on a priority basis. Since scheduling is performed daily, the number of jobs simultaneously considered is less than in the weekly system; thus the possibilities for evaluating priority interactions are fewer. To prevent a low priority job from remaining in the pool for an extremely long time, it might be beneficial to allow the priorities in the pool to "age" so that ultimately even a job which was initially very low in priority will be scheduled in preference to jobs which entered the pool much later with a higher priority. Other modifications of the priority scheme should be considered as well. For example, as a job approaches its due date, its priority might be automatically increased.

3. Timing

Because of the initial screening, some jobs will run on the same day they are submitted. Normally a job will be scheduled at the end of the day on which it was submitted. The scheduled time for such a job will be on day two, three, or four. Hence the leadtime is normally between one and three days. In periods of high loads, some jobs will remain in the pool until a later scheduling cycle.

A user can request a run time farther than three days away. His job request will be pooled but not scheduled until three days before the requested time.

4. Stability

The schedule once produced is relatively stable and will not change except when a pre-emptive job arrives or equipment fails.

5. Effect of Cancellation

When a job is cancelled, it will typically return to the pool for rescheduling. If it is desired to compensate this user for the inconvenience of cancellation, this can be done by increasing the priority of his job automatically when he returns to the pool.

6. Effect on the User

This system allows a job to be submitted and run on the same day if resources are available. The system can make use of user expressed preferences with respect to desired run times. Since most job requests are pooled until the end of the day, there is little opportunity for user interaction in the scheduling process.

7. Computer Facilities Required

This system if automated would require a terminal (or terminals) to be available at all times to accept requests and to do initial screening of new requests against the existing schedule. In addition a scheduling program would be run once a day.

8. Advantage-Disadvantage

The primary advantage of this system is in its ability to perform a reasonable degree of schedule optimization while operating in an almost continuous fashion. Its principal disadvantage is in the uncertain time delays between submission, scheduling and running of jobs because no schedule exists beyond the fourth day.

F. Summary

In Table 1 we bring together for convenient comparison the major results of the preceding analyses. It should be clear from the analyses and from this table that there are tradeoffs to be made in scheduling system design.

System	A. Weekly Batch	B. Continuous	C. Batch/ Continuous	D. Continuous, Initially Tentative	E. Daily Scheduling With 3-day Horizon
1. Schdule Optimization	Greatest chance for optimization Hardest schedul-ing alogrithm to design.	No chance for optimization. Scheduling logic is first-come-first-served.		Limited optimization in tentative period.	Substantial chance for optimization.
2. Priorities	Greatest chance for priority interactions.	No use of priorities in initial schedul-ing.		Limited priority interactions in the tentative period.	Substantial priority interactions.
3. Timing	High variance in lead times.	User has greatest control over tim-ing unless system becomes congested.	Mix between weekly batch and continuous depending on proportion of facilities allocated to each made.	Immediate repsonse, variable delay till run time.	Daily response, lead time usually 0-3 days.
4. Stability	Stable	Stable		Unstable beyond three day horizon. Tentative jobs frequently rearran-ged.	Stable
5. Effect of Cancellations	Rescheduling may have long lead time.	Immediately rescheduled, lack of optimization.		Immediately rescheduled. Wait variable	Re-enter pool for rescheduling at end of day.

Table I - Summary of Analyses

(Table continued)

Table I - Summary of Analyses (continued)

System	A. Weekly Batch	B. Continuous	C. Batch/ Continuous	D. Continuous Initially Tentative	E. Daily Scheduling With 3-day Horizon
6. Effects on the User	Easy to use and understand. No chance for interaction.	Easiest to use. Interaction possible.		Inability to plan in tentative period. Interaction only in initial scheduling which is tentative.	Easy to use. No chance for interaction.
7. Computer Facilities Required	One major scheduling run per week	Continuously available but minimal computation		Continuously available. More computation than continuous system.	One scheduling run per day.
8. Major Advantage	Possibility for schedule optimization	Easy to implement Immediate response		Immediate response and some schedule optimization.	Combines continuous operation with substantial optimization

Table I - Summary of Analyses

VI. CONCLUSIONS AND RECOMMENDATIONS

A. Conclusions

1. Several benefits can be obtained from an automated scheduling system. These include:

   - twenty-four-hour availability
   - more efficient schedules are possible
   - continuity in the scheduling process since it is not dependent on the skills of an individual scheduler.

2. Alternative systems can be compared by evaluating them in the following areas:

   - schedule efficiency, the extent to which the system allows schedule optimization to be performed
   - user convenience
   - the utilization of priority information
   - timing, the scheduling and running delays inherent in the system design
   - schedule stability; the extent to which a job, once scheduled is protected from being rescheduled
   - effect of job cancellations
   - computer facilities required.

3. Five alternative system designs are presented and evaluated in this report. Numerous variations are possible within each design, but either A (weekly batch) or E (continuous with 3 day horizon) appears to be most appropriate for FCDSSA's needs.

4. Whatever system design is selected, many details remain to be resolved including

   - the meaning and assignment of priorities
   - decision logic within the scheduling alogorithm

- communication interface with the users
- interfaces with management
- accurate determination of the computer requirements

B. Recommendations

1. FCDSSA continue with plans to implement an improved scheduling system.

2. FCDSSA examine the design alternatives presented here to determine which of the proposed systems is most appropriate for their needs.

3. FCDSSA consider their needs for a priority system and determine if the proposed priority structure consisting of maintenance, training, pre-emptive priority and regular priority jobs is adequate. Determine the number of regular priority classes and the procedures for assigning a pre-emptive priority.

4. Establish a timetable for the development and implementation of an improved scheduling system giving particular attention to the organizational implications of the transition. The user's needs must remain as the primary concern of the scheduling system itself, but the impacts on the scheduling department and maintenance must also be considered.